PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Protocol for a multicentre retrospective observational cohort study in Denmark Association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients
AUTHORS	Agerskov, Marianne; Thusholdt, Anna; Højlund, Jakob; Meyhoff, Christian; Sorensen, Henrik; Wiberg, Sebastian; Secher, Niels; Bang Foss, Nicolai

VERSION 1 – REVIEW

REVIEWER	Dr. George Kelley West Virginia University USA
REVIEW RETURNED	04-Jun-2019

GENERAL COMMENTS	GENERAL COMMENTS Thank you for the opportunity to review this protocol for a prospective observational cohort study examining the association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients. While this appears to be a worthwhile project, I have several suggestions for improvement, particularly with respect to the statistical handling and interpretation of data. There is also a need for someone with expertise in writing scientific articles in the English-language to rewrite some of this. Finally, there is a need to follow the journal's guidelines when revising this. SPECIFIC COMMENTS * Page 1 (Title) – I may be missing something here but this appears to be a retrospective cohort study given that the surgeries have already been conducted. Thus, here and throughout the rest of the manuscript, I would suggest that you refer to this study only as a retrospective cohort study. * Page 2 – There is no need for page 2. Please delete. * Page 3 (Abstract) – The abstract provides a balanced summary of what the authors plan to do. * Page 3 (Strengths and Limitations) – Please identify which of these are your strengths and which of these are your limitations. * Pages 4 and 5 (Introduction) – The scientific background and rationale for the proposed project is appropriate. In addition, both the objectives and pre-specified hypotheses are adequate.

	* Page 5 (Study Design) – Please provide a rationale for limiting this to major abdominal or hip fracture surgery here or in the next section. * Pages 5 and 6 (Participants, Inclusion/Exclusion Criteria) – With the exception of my comment above, the description of the participants, including the inclusion and exclusion criteria, are appropriate. * Page 6 (Data Collection) – The description of the plan for data collection is appropriate. * Page 6 (Exposure variables) – First, please provide a rationale for using 1 minute averages and the lowest 1 and 5 minute values for PPI, MAP and HR. Second, please provide data on the equipment (make/model/version, etc.) used to assess these variables, the validity and reliability of these instruments for assessing such, as well as any regular calibration procedures for this equipment. * Page 8 (Outcome Measures) – Why 30 and 90 days? Also, in Table 2, what about coding for postoperative complications after hospital discharge? * Pages 8 and 9 (Other Exposures) – What about patient characteristics such as age, gender, cigarette smoking, drug use (legal and illegal), body mass index (BMI), physical activity levels, etc.? * Pages 9 through 10 (3.10 Statistical Analysis) – Broadly, I would suggest that you consider restructuring your analysis based on the recent criticisms regarding use of the term 'statistically significant' and 'p<0.05'. An entire March issue of The American Statistician is devoted to this topic. The first article of the issue provides a nice overview of the articles in the issue as well as the author's opinions (see: Wasserstein, R. L., et al. (2019). "Moving to a world beyond "p < 0.05"." The American Statistician 73(sup1): 1-19). A related article also appears in a March issue of Nature (see: Amrhein, V., et al. (2019). "Scientists rise up against statistical significance." Nature 567: 305-307.) Along those lines, please tell the reader how you will consider not only statistical importance but also practical importance. I would also suggest that you calculate and report 95% confidence intervals, a metric that is now suggested by some to be called 'compatibility intervals' (see again: Amrhein, V., et al. (2019). "Scientists rise up against statistical significance." Nature 567: 305-307). To enhance interpretability, you may want to include the s-value, a statistic that is derived from your exact p-value (see: Greenland, S. (2019). "Valid p-values behave exactly as they should: Some misleading criticisms of p-values and their resolution with s-values." The American Statistician 73(sup1): 106-114). In the third paragraph and last sentence of the statistical analysis section, you talk about including significant exposure variables from univariate analyses in a multivariable logistic regression model. Along those lines, how do you plan on dealing with multi-collinearity between predictor variables? Also, you may want to consider calculating the recently developed E-value as a form of

	sensitivity analysis to account for unknown and/or unmeasured confounders (see: VanderWeele, T. J. and P. Ding (2017). "Sensitivity analysis in observational research: Introducing the E-value." Annals of Internal Medicine 167(4): 268-274). In the last paragraph, you talk about bootstrapping and say you will use 1,000 iterations. I would suggest that you run 10,000 versus 1,000 iterations or provide a reference that supports the use of 1,000 iterations for providing adequate coverage probabilities. Also, please tell the reader whether you will use parametric and/or non-parametric bootstrap resampling. Finally, please tell the reader how you plan on dealing with missing data for some variables and patients. * Page 10 (3.11 Patient and Public Involvement Statement) – While this is a protocol, this statement may need to be revised to reflect data on the patients from which this data is derived. * Page 12 (Availability of Data and Material) – Since this is a protocol, I would suggest that you revise this and tell the reader how data 'will' be made available.

REVIEWER	Thomas Weiser Stanford University
REVIEW RETURNED	26-Jun-2019

GENERAL COMMENTS	Agerskov and colleagues present the methods of a study they are performing to assess the utility of a peripheral perfusion index (PPI) in predicting postoperative morbidity and mortality in acute noncardiac surgery. They hypothesize that PPI will be more useful than other, traditional measures of hemodynamic status of surgical patients undergoing anesthesia, such as mean arterial pressure and indirect measurements of cardiac output. The study is well presented and the protocol is clear. The inclusion criteria are clear. The statistical analysis plan is somewhat concerning as I cannot see how they will use trends in PPI to help with determination. PPI is best used as an analysis of trend rather than isolation, so I do wonder how this will be assessed during the analysis. The capture of PPI is unclear, as is the device to be used. I assume PPI is continuously captured by the pulse oximeter device itself? If so, this should be stated. The calculation of PPI is noted as follows: "According to clinical routine, the staff measure MAP, HR, SAT, and thus PPI, and temperature (Temp) during surgery." (p6 line 33) However the actual calculation is not explained, or why from capturing MAP, HR, and SAT, the capture of PPI is understood. PPI "reflects the ratio between the pulsatile and non-pulsatile component of the arterial waveform as assessed by light traversing the tissue addressed, most often the finger, and it decreases in response to hypoperfusion" (p4 line 46) – presumably the Massimo device will be used? If not, how is PPI calculated?

	It is unclear how the researchers will account for the location of the oximetry device. Different PPIs are noted based on different sites (e.g. finger, ear, toe, etc). How will this be controlled for? The researchers propose to compare PPI to the more traditional parameters they are also collecting, which is encouraging. They will also assess the additional predictive value of PPI on the model, which is also good. They have not explained how they will deal with “noise” from the pulse oximeter. These devices can be quite variable in their ability to pick up the saturation signal, particularly in hemodynamically unstable patients with poor peripheral perfusion; how is this variability to be assessed?

REVIEWER	Dr Mark Edwards University Hospital Southampton and University of Southampton, UK I have received an honorarium for a lecture from Edwards Lifesciences (who are involved in haemodynamic monitoring) and am deputy Chief Investigator of the OPTIMISE II trial (Edwards Lifesciences and NIHR funded) although I do not receive financial support in this role. I am Chief Investigator for the FLO-ELA trial which is NIHR funded and supported by Deltex, Edwards Lifesciences and LiDCO, all of whom are involved in haemodynamic monitoring. I do not consider that these are competing interests in relation to this manuscript.
REVIEW RETURNED	26-Jul-2019

GENERAL COMMENTS	I thank the authors for the opportunity to review this protocol. Patients undergoing emergency surgery continue to suffer poor outcomes and we must continue to explore strategies that may improve their care. Optimal perioperative haemodynamic management in this group is yet to be defined but could have an important impact on outcomes. The authors propose an observational study to examine associations between a non-invasive measure of peripheral perfusion and clinical outcomes in acute hip fracture and abdominal surgery. I hope the suggestions below may be useful in strengthening this proposal. Key points: 1. As I understand it this study will examine data on patients that have already undergone surgery and had data collected about them on routine hospital systems at that time. I thought it was confusing to state that “we will collect data prospectively” (section 3.1 – and similar wording in section 3.4). I do not think the authors should define this as a “prospective” study. Unless I have misunderstood, this is a retrospective analysis of prospectively collected clinical data for surgical episodes already completed, so would be more accurately described as a retrospective cohort study. It is right that the authors only intend to collate and analyse these data after pre-specifying their methods and analysis plan but this is different from a true prospectively recruited patient cohort. 2. The two distinct surgical populations (hip fracture and abdominal surgery) may have very different clinical characteristics. The authors suggest they will analyse all patients together before performing subgroup analysis by surgery type. I wonder if separate primary analyses for each population would give a clearer result?

	3. It would be useful to know which technology had been used in these cases to derive the values for PPI 4. It would be helpful to list in much more detail (perhaps as an appendix) the planned outcome and exposure data points, for example: a) Baseline demographics – to include age, sex, pre-specified comorbidities or just composite Charlson? b) Intra-operative events – additional data fields capturing total fluid administered and type, the use of cardiac output monitoring and/or the use of protocolised fluid (?goal-directed) protocols would give useful context. c) Complications – although severity ranking using Clavien-Dindo is described, the authors do not describe how the occurrence of a complication will be defined (e.g. how will "postoperative infection" be defined?). This would be an important measure to guard against bias. 5. With reference to the STROBE checklist I am not sure why the following are listed as "not applicable" as they could be addressed in a protocol paper: a) 10 – study size justification. This is particularly important, particularly to demonstrate that there should be adequate outcome data points in a multivariable regression model. The authors have only described the time period for index surgical episodes but not discussed participant numbers. b) 12c and d – how missing data and loss to follow up will be dealt with (important in a retrospective analysis) 6. The protocol would be stronger if plans to minimise bias were described. For example, will data collectors viewing data on outcomes be blinded to PPI and other exposure variables? How will the morbidity outcomes be adjudicated? Minor points: 1. Introduction, line 39 "haemodynamic stability" – should this read "haemodynamic instability"? 2. Although as a retrospective cohort study there is no "participant experience" the authors could justify why patient and public involvement was not considered necessary. 3. 30- and 90-day mortality data will be apparently obtained from the patient electronic record. Can the authors confirm that deaths outside of hospital are reliably recorded in patient secondary care (hospital) records – or do all patients have an integrated single record with details of community and hospital care and vital status?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Dr. George Kelley
Institution and Country: West Virginia University USA

Reviewer:

Thank you for the opportunity to review this protocol for a prospective observational cohort study examining the association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients. While this appears to be a worthwhile project, I have several suggestions for improvement, particularly with respect to the statistical handling and interpretation of data. There is also a need for someone with expertise in writing scientific articles in the English-language to rewrite some of this. Finally, there is a need to follow the journal's guidelines when revising this.

Our response:

We are sincerely grateful to have received these insightful comments that we believe improve the protocol manuscript. Please find our point-by-point answers to each comment.

Reviewer:

Page 1 (Title) – I may be missing something here but this appears to be a retrospective cohort study given that the surgeries have already been conducted. Thus, here and throughout the rest of the manuscript, I would suggest that you refer to this study only as a retrospective cohort study.

Our response:

We agree, the intervention/the surgeries have already been conducted, therefore the study design is indeed retrospective. However, data collection is prospective i.e. after definition of hypothesis and primary/secondary endpoints, which we believe reduces bias. We have corrected this in the title and throughout the manuscript.

Reviewer:

Page 2 – There is no need for page 2. Please delete.

Our response:

Page 2 is deleted in main document_marked copy.

Reviewer:

Page 3 (Abstract) – The abstract provides a balanced summary of what the authors plan to do.

Our response: *Thank you.*

Reviewer:

Page 3 (Strengths and Limitations) – Please identify which of these are your strengths and which of these are your limitations

Our response:

These bullet strengths and limitations are required by the journal's guidelines and the layout is, to the best of our knowledge, required that way.

Reviewer:

Pages 4 and 5 (Introduction) – The scientific background and rationale for the proposed project is appropriate. In addition, both the objectives and pre-specified hypotheses are adequate.
Our response: *Thank you.*

Reviewer:

Page 5 (Study Design) – Please provide a rationale for limiting this to major abdominal or hip fracture surgery here or in the next section.
Our response:

Thank you for pointing this out. We believe that we provide rationale for the selection of study participants in the background section. Patients undergoing acute orthopaedic or abdominal surgery have high rates of postoperative complications and mortality.[1] *The morbidity and mortality are potentially associated with perioperative disturbances in macro and microvascular changes. We chose this group of patients as the most likely candidates for investigating the association between the PPI and morbidity, since investigation of patients with lower morbidity would require an even larger cohort, which would be beyond our logistical capacity.*

Reviewer:

Pages 5 and 6 (Participants, Inclusion/Exclusion Criteria) – With the exception of my comment above, the description of the participants, including the inclusion and exclusion criteria, are appropriate.
Our response: *Thank you.*

Reviewer:

Page 6 (Data Collection) – The description of the plan for data collection is appropriate.

Our response: *Thank you.*

1. Stoneham M, Murray D, Foss N: Emergency surgery: the big three--abdominal aortic aneurysm, laparotomy and hip fracture. Anaesthesia 2014; 69 Suppl 1:70–80

Reviewer:

Page 6 (Exposure variables) – First, please provide a rationale for using 1 minute averages and the lowest 1 and 5 minute values for PPI, MAP and HR. Second, please provide data on the equipment (make/model/version, etc.) used to assess these variables, the validity and reliability of these instruments for assessing such, as well as any regular calibration procedures for this equipment.

Our response:

Thank you for drawing our attention to this need for clarification. We have provided data on equipment and rationale for the chosen averages in the manuscript, marked in red, page 5

Reviewer:

Page 8 (Outcome Measures) – Why 30 and 90 days? Also, in Table 2, what about coding for postoperative complications after hospital discharge?

Our response:

We believe that 30- and 90-days mortality are widespread and accepted outcome measures and chosen as such. We obtain postoperative complications, classified by Clavien-Dindo Classification of Surgical Complications 30 days after surgery. All information regarding complications after discharge as well as information on date and cause of death will be available in the electronic medical chart.

Reviewer:

Pages 8 and 9 (Other Exposures) – What about patient characteristics such as age, gender, cigarette smoking, drug use (legal and illegal), body mass index (BMI), physical activity levels, etc.?
Our response:

Patient demographics, including age and gender will be obtained from the medical record as disclosed. In order to assess and classify comorbidity and physical status we use 3 accepted classification systems; American Society of Anesthesiology (ASA), WHO/ECOG/Zubrod score and The Charlson Comorbidity Index, also disclosed in the manuscript.

Reviewer:

Pages 9 through 10 (3.10 Statistical Analysis) – Broadly, I would suggest that you consider restructuring your analysis based on the recent criticisms regarding use of the term ‘statistically significant’ and ‘p<0.05’. An entire March issue of The American Statistician is devoted to this topic. The first article of the issue provides a nice overview of the articles in the issue as well as the author’s opinions (see: Wasserstein, R. L., et al. (2019). "Moving to a world beyond “p < 0.05”." The American Statistician 73(sup1): 1-19). A related article also appears in a March issue of Nature (see: Amrhein, V., et al. (2019). "Scientists rise up against statistical significance." Nature 567: 305-307.) Along those lines, please tell the reader how you will consider not only statistical importance but also practical importance. I would also suggest that you calculate and report 95% confidence intervals, a metric that is now suggested by some to be called ‘compatibility intervals’ (see again: Amrhein, V., et al. (2019). "Scientists rise up against statistical significance." Nature 567: 305-307).

To enhance interpretability, you may want to include the s-value, a statistic that is derived from your exact p-value (see: Greenland, S. (2019). "Valid p-values behave exactly as they should: Some misleading criticisms of p-values and their resolution with s-values." The American Statistician 73(sup1): 106-114).

In the third paragraph and last sentence of the statistical analysis section, you talk about including significant exposure variables from univariate analyses in a multivariable logistic regression model. Along those lines, how do you plan on dealing with multi-collinearity between predictor variables? Also, you may want to consider calculating the recently developed E-value as a form of sensitivity analysis to account for unknown and/or unmeasured confounders (see: VanderWeele, T. J. and P. Ding (2017). "Sensitivity analysis in observational research: Introducing the E-value." Annals of Internal Medicine 167(4): 268-274).

In the last paragraph, you talk about bootstrapping and say you will use 1,000 iterations. I would suggest that you run 10,000 versus 1,000 iterations or provide a reference that supports the use of 1,000 iterations for providing adequate coverage probabilities. Also, please tell the reader whether you will use parametric and/or non-parametric bootstrap resampling.

Finally, please tell the reader how you plan on dealing with missing data for some variables and patients.

Our response:

Thank you very much for providing us with a very thorough and knowledgeable review on our statistical analysis plan. We are aware of, and often discuss, the use of the widely used and accepted term "statistically significant" and the somewhat "blind use" of p-values. However, the "traditional" way of analysing epidemiological data is, in our opinion, still the most widespread and accepted way, and may therefore be more understandable and applicable when interpreting results in a clinical context.

We plan to perform multivariate logistic analysis to assess the associations between peripheral perfusion and outcome, not only providing p-values, but also reporting OR's and 95% confidence intervals. We plan to carefully discuss the clinical significance and importance of our findings.

We plan to test clinically and physiologically relevant secondary exposure variables in univariate analysis and then include these in the multivariate model in case of statistical significance. With regards to multi-collinearity, we plan to include interaction links between MAP and PPI as specified in the protocol.

Thank you for guiding our attention to the lack in describing our plan on dealing with missing data. If missing data on exposure variables exceeds 10%, we plan to perform multiple imputation and if a large fraction of data is imputed, we wish to compare observed and imputed values. If missing data on outcome variables exceeds 10%, we plan to manually impute worst/best case scenarios and perform subgroup analysis. The manuscript has been revised accordingly.

Reviewer:

Page 10 (3.11 Patient and Public Involvement Statement) – While this is a protocol, this statement may need to be revised to reflect data on the patients from which this data is derived.

Our response:

Thank you for pointing this out for us. We argue that this is a retrospective observational study with no involvement of patients. Findings may have immediate impact on clinical practice with no inconvenience to future patients.

Reviewer:

Page 12 (Availability of Data and Material) – Since this is a protocol, I would suggest that you revise this and tell the reader how data 'will' be made available.

Our response:

We have revised the manuscript as suggested.

Reviewer: 2

Reviewer Name: Thomas Weiser
Institution and Country: Stanford University

Reviewer:

The study is well presented and the protocol is clear.
Our response:

We are sincerely grateful to have received these insightful comments that we believe improve the protocol manuscript. Please find our point-by-point answers to each comment.

Reviewer:

The inclusion criteria are clear.
Our response: *Thank you.*

Reviewer:

The statistical analysis plan is somewhat concerning as I cannot see how they will use trends in PPI to help with determination. PPI is best used as an analysis of trend rather than isolation, so I do wonder how this will be assessed during the analysis.
Our response:

Thank you for the comments. In the intraoperative patient charts (EPIC), PPI is presented as columns of averages, generated on a time interval. Data is lifted sequentially at different time intervals. This is not the same as an analysis of rolling averages, and we have noted this restriction in the methods section.

Reviewer:

The capture of PPI is unclear, as is the device to be used. I assume PPI is continuously captured by the pulse oximeter device itself? If so, this should be stated. The calculation of PPI is noted as follows: "According to clinical routine, the staff measure MAP, HR, SAT, and thus PPI, and temperature (Temp) during surgery." (p6 line 33) However the actual calculation is not explained, or why from capturing MAP, HR, and SAT, the capture of PPI is understood. PPI "reflects the ratio between the pulsatile and non-pulsatile component of the arterial waveform as assessed by light

traversing the tissue addressed, most often the finger, and it decreases in response to hypoperfusion" (p4 line 46) – presumably the Massimo device will be used? If not, how is PPI calculated?
Our response:

Thank you for drawing attention to this need for clarification. We have revised the manuscript accordingly.

Reviewer:

It is unclear how the researchers will account for the location of the oximetry device. Different PPIs are noted based on different sites (e.g. finger, ear, toe, etc). How will this be controlled for?
Our response:

As we obtain retrospective data, we have no influence on the location of the oximetry device. The location of the probe is not registered in the medical chart although it is most often, by routine, placed on the second or third finger.

Reviewer:

The researchers propose to compare PPI to the more traditional parameters they are also collecting, which is encouraging. They will also assess the additional predictive value of PPI on the model, which is also good.
Our response:

Thank you, we are additionally encouraged by the reviewer's comment.

Reviewer:

They have not explained how they will deal with "noise" from the pulse oximeter. These devices can be quite variable in their ability to pick up the saturation signal, particularly in hemodynamically unstable patients with poor peripheral perfusion; how is this variability to be assessed?
Our response:

Artefactual data is a challenge when dealing with large automatically obtained datasets.

We thoroughly investigated the possibility of exporting perioperative data from the electronic anaesthesia chart to our database, however this was not feasible. Instead, we manually transfer data, as disclosed in the manuscript. We acknowledge that this method increases the risk of typing errors, but also argue that we have the possibility of identifying artefactual data.

Reviewer: 3

Reviewer Name: Dr Mark Edwards
Institution and Country: University Hospital Southampton and University of Southampton, UK
Competing interests: I have received an honorarium for a lecture from Edwards Lifesciences (who are involved in haemodynamic monitoring) and am deputy Chief Investigator of the OPTIMISE II trial (Edwards Lifesciences and NIHR funded) although I do not receive financial support in this role. I am

Chief Investigator for the FLO-ELA trial which is NIHR funded and supported by Deltex, Edwards Lifesciences and LiDCO, all of whom are involved in haemodynamic monitoring. I do not consider that these are competing interests in relation to this manuscript.

Reviewer:
As I understand it this study will examine data on patients that have already undergone surgery and had data collected about them on routine hospital systems at that time. I thought it was confusing to state that "we will collect data prospectively" (section 3.1 – and similar wording in section 3.4). I do not think the authors should define this as a "prospective" study. Unless I have misunderstood, this is a retrospective analysis of prospectively collected clinical data for surgical episodes already completed, so would be more accurately described as a retrospective cohort study. It is right that the authors only intend to collate and analyse these data after pre-specifying their methods and analysis plan but this is different from a true prospectively recruited patient cohort.
Our response:

We are very honoured to have been given the opportunity to revise our protocol manuscript according to these insightful and valuable reviewer comments.

The study is indeed of retrospective data and the manuscript is revised accordingly.

Reviewer:

The two distinct surgical populations (hip fracture and abdominal surgery) may have very different clinical characteristics. The authors suggest they will analyse all patients together before performing subgroup analysis by surgery type. I wonder if separate primary analyses for each population would give a clearer result?

Our response:

This is a very valid point and the reason we have preplanned a subanalysis of the individual populations. The impact of different perioperative physiological disturbances such as vasoplegia, vasoconstriction, hypovolaemia, septicaemia or anaemia on microcirculation is poorly understood. We believe – unpublished data in analysis - that sympatholysis is the main determinant of the association between PPI and central organ perfusion, making the analysis of hip fracture undergoing general anaesthesia together with emergency laparotomy potentially more relevant that the individual pathologies analyzed individually. Accordingly, we believe that an overall analysis followed by secondary analysis on pathology and anaesthetic technique is appropriate.

Reviewer:

It would be useful to know which technology had been used in these cases to derive the values for PPI

Our response:

Thank you for drawing our attention to this. We have now provided information in the manuscript.

Reviewer:

It would be helpful to list in much more detail (perhaps as an appendix) the planned outcome and exposure data points, for example:

a) Baseline demographics – to include age, sex, pre-specified comorbidities or just composite Charlson?

Our response: *We believe that we provide information in Table 1, 2 and 3 on the planned exposure and outcome datapoints. We agree that the statement "Baseline demographics" might be inadequate. In order to assess and classify comorbidity and physical status we use 3 accepted classification systems; American Society of Anesthesiology (ASA), WHO/ECOG/Zubrod score and The Charlson Comorbidity Index as disclosed in the manuscript.*

b) Intra-operative events – additional data fields capturing total fluid administered and type, the use of cardiac output monitoring and/or the use of protocolised fluid (?goal-directed) protocols would give useful context.

Our response: *We believe that we try to provide many of the mentioned additional data fields for context. We describe (Table1) the intraoperative values we plan to obtain including Vasoactive medication, Anaesthetic method, total operation time ect. We have planned, and designed the database accordingly, to obtain total fluid administered, the manuscript has been revised accordingly. We plan to carefully describe the setting when presenting the results, as some of the patients will have received protocolised fluids and some will not, but due to the retrospective nature of this study it will not be possible to standardize this in a meaningful non-biased way when analysing data. We will of course deal with this when discussing the results.*

c) Complications – although severity ranking using Clavien-Dindo is described, the authors do not describe how the occurrence of a complication will be defined (e.g. how will "postoperative infection" be defined?). This would be an important measure to guard against bias.

Our response: *The Clavien-Dindo Classification of Surgical Complications provides instructions for severity grading and is based on the treatment of complications, not on the complication in itself, i.e: Gade II "requiring pharmacological treatment", grade IIIa: "Surgical, endoscopic, or radiological intervention". We review the medical charts and classify accordingly whenever patients have complications that need treatment. We believe it reduces bias, since we do not define or interpret when patients might have had a complication, i.e infection, but we simply register treatment.*

Reviewer:

With reference to the STROBE checklist I am not sure why the following are listed as "not applicable" as they could be addressed in a protocol paper:
a) 10 – study size justification. This is particularly important, particularly to demonstrate that there should be adequate outcome data points in a multivariable regression model. The authors have only described the time period for index surgical episodes but not discussed participant numbers.
b) 12c and d – how missing data and loss to follow up will be dealt with (important in a retrospective analysis)

Our response:

We acknowledge these points as very important when handling observational data.

a) *The estimated probability of any severe complication or death within 30 days for high risk acute abdominal or hip fracture surgery is approximately 45% and 30% in our setting respectively. (ref in text) A reasonable estimation of overall severe postoperative complications or death within 30 days is pragmatically set at 40%. We expect to include around 2300 patients in the cohort with approximately equal distribution between class of surgery. We plan to perform logistic regression models evaluating the association between the primary outcome and PPI. Using Whittemore´s formula (ref in text) requiring a dichotomous dependent variable and continuous risk factors, a sample size of 2300 patients with an event rate of 40% will enable us to detect an Odds Ratio of 0.9 using a two-tailed test with a significance level of 5% and a power of 80%.*
b) *If missing data on exposure variables exceeds 10%, we plan to perform multiple imputation as a secondary analysis. We will compare observed and imputed values. If missing data on*

outcome variables exceeds 10%, we plan to manually impute worst/best case scenarios and perform subgroup analysis. The manuscript has been revised accordingly.

Reviewer:

The protocol would be stronger if plans to minimise bias were described. For example, will data collectors viewing data on outcomes be blinded to PPI and other exposure variables? How will the morbidity outcomes be adjudicated?

Our response:

Due to limited resources and the rather laborious nature of the data collection process we recognize the risk of bias. Double entry of data will not be logistically possible if the data is to be collected within the financial constraints of the project. The same person entering hemodynamic variables from the anaesthesia chart to the database will be reviewing the medical chart and categorizing postoperative complications according to clavien dindo. However, we argue that due to the nature of the clavien dindo classification, accounted for earlier, the risk of bias is reduced.

Data entry personnel will obtain data on both orthopaedic and abdominal surgery patients from both hospitals minimizing the risk of systematic bias. Whenever questions arise there will be conference between data entry personnel.

We now discuss this in the protocol limitations and will address this in more detail when discussing results.

Introduction, line 39 “haemodynamic stability” – should this read “haemodynamic instability”?

Our response: *Thank you, we have corrected this in the manuscript.*

Reviewer:

Although as a retrospective cohort study there is no “participant experience” the authors could justify why patient and public involvement was not considered necessary.
Our response:

We choose not to include this in the protocol as all data will be retrieved from medical records and anaesthesia charts without any impact on future treatment for the involved patients. The study will be carried out and is approved as a quality approving study in the involved departments.

Reviewer:

30- and 90-day mortality data will be apparently obtained from the patient electronic record. Can the authors confirm that deaths outside of hospital are reliably recorded in patient secondary care (hospital) records – or do all patients have an integrated single record with details of community and hospital care and vital status?

Our response:

We confirm that if patients die out of hospital, information on date of death will appear in the medical chart for all patients with a Danish civil registration number.

VERSION 2 – REVIEW

REVIEWER	Dr. George Kelley West Virginia University, USA
REVIEW RETURNED	05-Sep-2019

GENERAL COMMENTS	GENERAL COMMENTS Thank you for the opportunity to review this revised protocol for a prospective observational cohort study examining the association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients. Overall, the authors have been responsive to my previous comments. I have only three remaining minor comments, listed below. SPECIFIC COMMENTS * Page 2 (Strengths and Limitations) – While I agree with your response to my previous comment about the journals requirements regarding strengths and limitations, the more naïve reader may have a difficult time in discerning between the two as it is currently written. Therefore, I would suggest that you expand each of these bulleted items and say why each is a strength or weakness. As an example, please see our previously published work in BMJ Open: Kelley, G. A., et al. (2018). "Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a systematic review with meta-analysis of randomised controlled trials." BMJ Open 8(2): 18. * Page 9, last three sentences of paragraph 3 – Per my previous comments, I would suggest that you state that you will run 10,000 versus 1,000 iterations for your bootstrap resampling or provide a reference that supports the use of 1,000 iterations for providing adequate coverage probabilities. Also, please tell the reader whether you will use parametric and/or non-parametric bootstrap resampling. * While I'm fine with your response about the issues I raised regarding the use of the term "statistical significance" given that it is still the traditional way that things are done, it's important to understand that nothing will change unless investigators start making an effort to do so. END OF REVIEW

REVIEWER	Dr Mark Edwards University Hospital Southampton and University of Southampton, UK I have received an honorarium for a lecture from Edwards Lifesciences

	(who are involved in haemodynamic monitoring) and am deputy Chief Investigator of the OPTIMISE II trial (Edwards Lifesciences and NIHR funded) although I do not receive financial support in this role. I am Chief Investigator for the FLO-ELA trial which is NIHR funded and supported by Deltex, Edwards Lifesciences and LiDCO, all of whom are involved in haemodynamic monitoring. I do not consider that these are competing interests in relation to this manuscript.
REVIEW RETURNED	25-Sep-2019

GENERAL COMMENTS	Many thanks for asking me to review this revised manuscript. I thank the authors for their thoughtful responses to my comments. The authors have made suitable adjustments to the manuscript – including discussing some unavoidable limitations where relevant – in response to the following areas I raised: -Analysing the two surgical categories together or separately -Technology used to derive PPI -Intraoperative event data collection -Sample size justification and missing data strategy -Source of mortality data However, there are a number of issues I raised which I feel still require some further revision to address fully: 1. Study description. Regarding the definition of the study (prospective vs. retrospective) the authors have appropriately revised the title of the study. However, it is still referred to as "prospective" in the Abstract (methods and analysis) and the "strengths and limitations of this study" section. In this latter section the study is described as "retrospective data, prospective data collection". As all source data has already been recorded within the medical record prior to the start of the study design, I think this is better described as "retrospective analysis of prospectively collected clinical data", and in summary as a "retrospective study" throughout (including "Study Design" section). Similarly in "Study Design" it should read "...an observational cohort study design of patients who underWENT acute major abdominal… 2. Baseline/demographic data. The authors have not clarified whether the most basic demographics (age, sex) will be collected and presented? 3. Outcomes definitions. The authors argue that the use of the Clavien-Dindo system alone is adequate to define the occurrence of a complication. However, I remain concerned that using this severity grading system alone without also including a formal, standardised definition of the occurrence of a complication (e.g. the StEP-COMPAC [Myles BJA 2016] or EPCO [Jammer Eur J Anaesth 2015] recommended end point definitions) does raise the possibility of differing interpretation by data collectors and therefore bias. This may also limit comparisons of this study with others and the ability for others to replicate this work. If the authors are committed to using only Clavien-Dindo, it may be best to discuss this potential limitation in this protocol manuscript and the ultimate output. 4. Patient involvement. I think the justification given for not involving patients/public by the authors is reasonable, and I would recommend that this justification is added to that section in the manuscript.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Dr. George Kelley
Institution and Country: West Virginia University, USA Please state any competing interests or state 'None declared': None declared.

Please leave your comments for the authors below GENERAL COMMENTS

Thank you for the opportunity to review this revised protocol for a prospective observational cohort study examining the association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients. Overall, the authors have been responsive to my previous comments. I have only three remaining minor comments, listed below.

Reviewer:
* Page 2 (Strengths and Limitations) – While I agree with your response to my previous comment about the journals requirements regarding strengths and limitations, the more naïve reader may have a difficult time in discerning between the two as it is currently written. Therefore, I would suggest that you expand each of these bulleted items and say why each is a strength or weakness. As an example, please see our previously published work in BMJ Open: Kelley, G. A., et al. (2018). "Community-deliverable exercise and anxiety in adults with arthritis and other rheumatic diseases: a systematic review with meta-analysis of randomised controlled trials." BMJ Open 8(2): 18.

Our response:
Thank you for further elaborating on this point and providing a useful example. We have expanded the bulleted items, hopefully aiding discrimination between strengths and limitations.

Reviewer:
* Page 9, last three sentences of paragraph 3 – Per my previous comments, I would suggest that you state that you will run 10,000 versus 1,000 iterations for your bootstrap resampling or provide a reference that supports the use of 1,000 iterations for providing adequate coverage probabilities. Also, please tell the reader whether you will use parametric and/or non-parametric bootstrap resampling.

Our response:
Thank you.
For internal validation we plan to use non-parametric bootstrap resampling, as we don't know the distribution of data beforehand. Upon request, we will run 10,000 iterations.

Reviewer:
* While I'm fine with your response about the issues I raised regarding the use of the term "statistical significance" given that it is still the traditional way that things are done, it's important to understand that nothing will change unless investigators start making an effort to do so.

Our response:
Thank you for once again pointing this out. We intend to carefully discus this when presenting results hopefully aiding focus on clinical significance rather than statistical significance.

Reviewer: 3

Reviewer Name: Dr Mark Edwards
Institution and Country: University Hospital Southampton and University of Southampton, UK
Please state any competing interests or state 'None declared':

I have received an honorarium for a lecture from Edwards Lifesciences (who are involved in haemodynamic monitoring) and am deputy Chief Investigator of the OPTIMISE II trial (Edwards Lifesciences and NIHR funded) although I do not receive financial support in this role. I am Chief Investigator for the FLO-ELA trial which is NIHR funded and supported by Deltex, Edwards Lifesciences and LiDCO, all of whom are involved in haemodynamic monitoring. I do not consider that these are competing interests in relation to this manuscript.

Reviewer:
Please leave your comments for the authors below Many thanks for asking me to review this revised manuscript. I thank the authors for their thoughtful responses to my comments. The authors have made suitable adjustments to the manuscript – including discussing some unavoidable limitations where relevant – in response to the following areas I raised:
-Analysing the two surgical categories together or separately -Technology used to derive PPI -Intraoperative event data collection -Sample size justification and missing data strategy -Source of mortality data
1.
However, there are a number of issues I raised which I feel still require some further revision to address fully:
Study description. Regarding the definition of the study (prospective vs. retrospective) the authors have appropriately revised the title of the study. However, it is still referred to as “prospective” in the Abstract (methods and analysis) and the “strengths and limitations of this study” section. In this latter section the study is described as “retrospective data, prospective data collection”. As all source data has already been recorded within the medical record prior to the start of the study design, I think this is better described as “retrospective analysis of prospectively collected clinical data”, and in summary as a “retrospective study” throughout (including “Study Design” section). Similarly in “Study Design” it should read “...an observational cohort study design of patients who underWENT acute major abdominal….

Our response:
Thank you for the thorough review, which should not have been necessary, we apologize. The manuscript has been revised accordingly.

Reviewer:
2.
Baseline/demographic data. The authors have not clarified whether the most basic demographics (age, sex) will be collected and presented?

Our response:
We collect basic demographic data (age, sex, height and weight), which is now introduced in the manuscript.

Reviewer:
3. Outcomes definitions. The authors argue that the use of the Clavien-Dindo system alone is adequate to define the occurrence of a complication. However, I remain concerned that using this severity grading system alone without also including a formal, standardised definition of the occurrence of a complication (e.g. the StEP-COMPAC [Myles BJA 2016] or EPCO [Jammer Eur J Anaesth 2015] recommended end point definitions) does raise the possibility of differing interpretation by data collectors and therefore bias. This may also limit comparisons of this study with others and the ability for others to replicate this work. If the authors are committed to using only Clavien-Dindo, it may be best to discuss this potential limitation in this protocol manuscript and the ultimate output.

Our response:
Thank you for further pointing our attention to this and providing very relevant and useful references. Amending our data collection set-up is, unfortunately, not a possibility. We agree that including a formal, standardized definition of the occurrence of a complication may limit possible bias from differing interpretation by data collectors. However, the Clavien Dindo severity grading system is included in StEP-COMPAC and the use has demonstrated high inter-individual agreement when assessing postoperative complications.
We briefly touch upon this in the “limitations and bias” section of the protocol and plan to further elaborate when presenting and discussing results.

Reviewer:
4. Patient involvement. I think the justification given for not involving patients/public by the authors is reasonable, and I would recommend that this justification is added to that section in the manuscript.

Our response:
Thank you. The mentioned justification has been entered in to the "Patient and Public involvement Statement)

VERSION 3 – REVIEW

REVIEWER	Dr. George Kelley West Virginia University, USA
REVIEW RETURNED	17-Oct-2019

GENERAL COMMENTS	Thank you for the opportunity to review this second revised protocol for a prospective observational cohort study examining the association between the intraoperative peripheral perfusion index and postoperative morbidity and mortality in acute non-cardiac surgical patients. The authors have been responsive to my previous three comments. The only remaining and very minor comment is that on the second to last line of the Strengths and Limitations section it appears that the word "lager" should be replaced with "large"?

REVIEWER	Dr Mark Edwards University Hospital Southampton and University of Southampton, UK As per previous reviews
REVIEW RETURNED	25-Oct-2019

GENERAL COMMENTS	Many thanks to the authors for further considering my comments. I have reviewed the updated manuscript and confirm that these have now all been suitably addressed. I have spotted a minor typographical error on p6 - strengths and limitations - "lager" should read "larger".